COMPARISON OF MACHINE LEARNING TECHNIQUES FOR INDIRECT ASSESSMENT METHODS OF BODY CORE TEMPERATURE

Ricardo Vardasca[1,2,3]

1 ISLA Santarem, Santarem, Portugal
2 INEGI, Porto, Portugal
3 Faculty of Computing, Engineering and Science, University of South Wales, Pontypridd, UK

Introduction

Pandemic conditions are once again in great prominence with the recent situation caused by COVID-19, some of these conditions present feverish states that can be detected by means of mass screening at places of great influx of people. There are available different indirect methods to estimate human body core temperature. Being a febrile state considered of a body core temperature higher than 37.5 ºC. This value may differ according to the indirect method used, which can make it difficult to identify febrile cases close to the threshold value, for assisting in this task advanced Artificial Intelligence tools such as Machine Learning (ML) algorithms may be an important aid. The aim of this research is to evaluate which ML technique has the best performance with a certain indirect method of assessing body temperature, considering the reference provided by another method.

Methods

Among the available methods to indirectly estimate human body core temperature are the axillar and tympanic thermometers and the infrared measurements of the forehead and the inner canthi of the eye.

A total of 140 subjects (37±7.1 years old, ranging from 21 to 63, 69 males and 71 females), from which 10 were considered febrile by the axilla thermometer assessment. All were screened with axillar thermometer (Beurer FT09/1 with precision of ± 0.1 °C, Ulm - Germany) and tympanic thermometer (Sanitas SFT 75 with precision of ± 0.2 °C, Uttenweiler - Germany) and facial infrared imaging using a thermal camera FLIR T430 (FPA array size of 320x240, NETD of <50mK @ 30ºC and measurement uncertainty of ±2% of the overall temperature reading) at a health center of Santarem.

To identify the febrile state, it was used the threshold value of 37.5ºC on the axillar thermometer reading as a reference, being the measurements taken in both armpits and considered the average when there was no difference above 0.3ºC between both sides, being the measurement repeated if the difference was higher. In the tympanic membrane thermometer, the two ear canals were assessed and any measurement with a bilateral difference major than 0.3º C

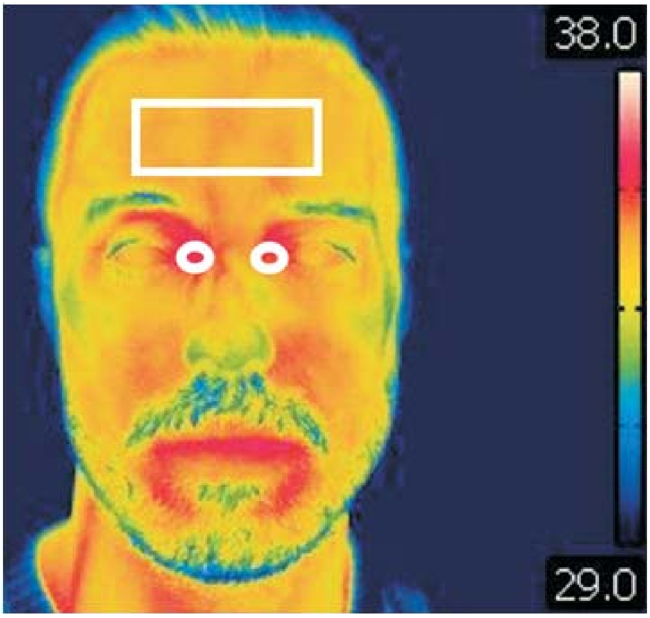

Figure 1
The ROIs at the frontal face thermal images.

was disregarded and the mean between sides was considered. The thermal images were collected using the standard recommendations on room, equipment, and participant preparation. The average room temperature was 22.1±0.2 °C and relative humidity o 47±1.8 %. In each thermogram three ROIs (Fig. 1) were considered, one in the forehead and others at each inner canthi of the eye, being the mean temperatures recorded. For the inner canthi of the eye the average of both sides was considered.

Five ML methods were selected for this research: Multilayer Perceptron (MLP), Support Vector Machines (SVM), Naïve Bayes (NB), k-Nearest Neighbor (kNN) and Random Forest (RF). A Python script was developed consisting of six tests: 5 and 10-fold cross section with 20%, 25% and 30% sample testing. When a 20% test size was used, it meant that the ML model was trained with the remaining 80% sample size. The use of several N-fold cross validation is important to vary the composition of the test and training sets that are selected randomly from the sample.

To verify which ML method would perform better and with which measurement method the overall accuracy (the

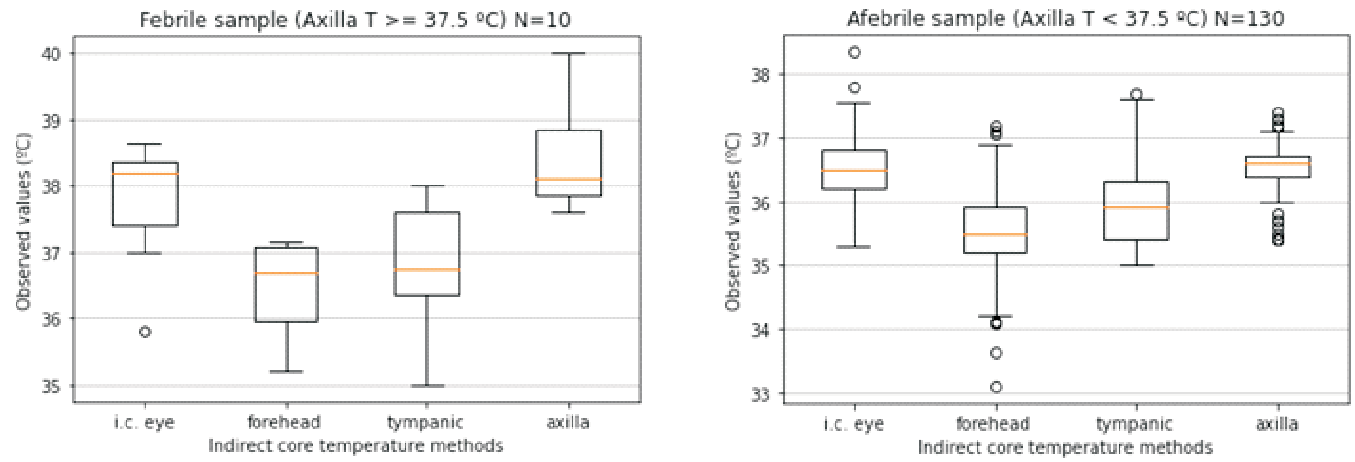

Figure 2
The data sample distribution per indirect core temperature method, left the febrile group and right the afebrile group (based on the axilla thermometer assessment).

ratio between number of correct predictions by the total number of predictions. Being calculated in this experimental work by the ratio of the sum of the true positives and true negatives by the sum of the true positives, false positives, true negatives, and false negatives) of each ML method was calculated.

Results

The sample data distribution of the recorded values with the different core temperature estimation methods is presented at figure 2. The results at the defined tests (sample testing size and fold cross validation) for all ML techniques and temperature assessment methods are showed in table 1, from which it can be observed that the worst results are given by MLP for all tests and methods. All the other ML techniques presented good results (accuracy > 85%), being the SVM the technique that presented better results for the inner canthi of the eye and forehead measurement, being closely followed by the NB algorithm. For tympanic measurements, the ML technique that overperformed was the NB followed by the RF.

Discussion and conclusions

Different indirect methods to assess body core temperature present different temperature values as showed by Vardasca et al., 2019. Therefore, different fever thresholds may have to be considered, which can make hard the life of assessors (humans or systems), to automatize the process ML techniques may be used to help the burden. This research showed that the ML algorithm performance may differ with the temperature assessment method, Artificial Neural Networks (MLP) based systems should be avoided since there is one unique value of entrance to find an output, techniques such as SVM, NB and RF presented better performance (accuracy > 89%) and are more reliable for this kind of implementation. The knowledge base size, for training the ML models, is also an aspect of importance since values may differ from having a 70 to 80% training sample. Although for the suggested methods this variation

Table 1
Accuracy of the Machine Learning techniques on the body temperature assessment methods within the different configuration tests.

Method	Test	5-fold cross validation			10-fold cross validation		
		20%	**25%**	**30%**	**20%**	**25%**	**30%**
I.C. Eye	MLP	63.00	40.00	40.66	50.00	45.20	49.66
	SVM	95.00	96.80	94.00	95.00	94.40	96.00
	NB	95.00	95.20	94.00	94.50	94.00	96.33
	kNN	94.00	96.00	92.00	95.00	90.40	92.00
	RF	92.00	97.60	92.66	95.00	92.00	93.66
Forehead	MLP	39.45	42.70	40.54	75.67	48.10	40.54
	SVM	91.89	93.51	92.97	91.89	95.40	91.08
	NB	91.89	94.05	91.89	91.89	94.86	93.24
	kNN	85.94	88.10	85.40	85.67	91.89	87.29
	RF	89.73	92.97	90.27	90.00	93.51	90.81
Tympanic	MLP	37.83	41.08	40.54	75.67	24.05	31.89
	SVM	90.27	92.97	92.97	94.05	91.08	94.86
	NB	91.89	93.51	93.51	95.40	91.35	96.21
	kNN	89.18	88.64	88.10	88.91	86.75	88.10

is smaller. Nonetheless this experiment should be validated in a larger population size for better understanding.

Acknowledgments

This research was partial funding by the projects LAETA-UIDB/50022/2020 and UIDP/50022/2020.

References

Ring EFJ, Jung A, Kalicki B, Zuber J, Rustecka A, Vardasca R. New standards for fever screening with thermal imaging systems. Journal of Mechanics in Medicine and Biology 2013; 13(02), 1350045.

Vardasca, R.; Magalhaes, C.; Marques, D.; Moreira, J.; Frade, R.; Seixas, A.; Mendes, J; Ring, F. Bilateral assessment of body core temperature through axillar, tympanic and inner canthi thermometers in a young population. Physiological measurement, 2019, volume40(9), 094001.

Vardasca Rn Vaz L, Mendes J. Classification and decision making of medical infrared thermal images. Classification in BioApps 2018. pp. 79-104.

Ammer K, Ring F. The Thermal Human Body: A Practical Guide to Thermal Imaging. CRC Press 2019.

RussellS; Norvig P. Artificial intelligence: a modern approach; 2020.

Corresponding author
Prof Ricardo Vardasca PhD
ISLA Santarem,
Largo Cândido dos Reis,
2000-241 Santarem,
Portugal
Email: ricardo.vardasca@islasantarem.pt

(Revision accepted 26.04.2021)